

University of Groningen

Ouderdom, omvang en citatiescores: rankings nader bekeken

van Rooij, Jules

Published in:
THEMA, Tijdschrift voor Hoger Onderwijs en Management

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2017

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

van Rooij, J. (2017). Ouderdom, omvang en citatiescores: rankings nader bekeken. THEMA, Tijdschrift voor Hoger Onderwijs en Management, 24(3), 77-83.

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

In een serie gaat de dataredactie van *Thema* in op de vraag of de in Nederland vigerende rankings wel de aandacht verdienen die ze krijgen. Voor deze tweede aflevering onderzoekt Jules van Rooij in hoeverre omvang en leeftijd van een instelling gecorreleerd zijn met de positie in een ranking en hoe dit zich verhoudt tot citatie-impactindicatoren.

Ouderdom, omvang en citatiescores

Rankings nader bekeken

Dataredactie, Jules van Rooij

In de eerste aflevering van deze reeks bleek dat vooral universiteiten in de Randstad hoog scoren in de mondiale reputatierankings. De totaalscore op de *Times Higher Education* (THE) en de *QS World University Ranking* (QS) correleren in belangrijke mate met de reputatiescores, en in het geheel niet met citatiescores, noch met studentenoordelen in de Nationale Studenten Enquête (NSE). Wij opperden dat zowel omvang als bekendheid van de instelling (en van de stad) daarbij een rol zou kunnen spelen. Beide lijken belangrijker dan onderzoekskwaliteit en citatie-impact.

Om dit vermoeden nader te onderzoeken, gaan we nu in op de vraag in hoeverre grootte en leeftijd van een instelling gecorreleerd zijn met de positie in een ranking en hoe dit zich verhoudt tot citatie-impactindicatoren. Daartoe hebben we onze analyse uitgebreid met buitenlandse universiteiten (alle in de vier Top-300-lijsten) en met twee andere wereldwijde rankings, de *Shanghai ARWU* en de *Leiden Ranking*. Deze laatste nemen geen reputatiescores op. De Leiden Ranking is bovendien niet zozeer een ranking als wel een online tool om universiteiten te vergelijken op een aantal zelf te kiezen indicatoren, ontleend aan de citatiedatabase *Web of Science*. De onderliggende methodologie¹ wordt breed erkend als degelijk en betrouwbaar – reden waarom de Leiden Ranking is meegenomen als ‘benchmark’ voor een aantal onderzoekindicatoren van de andere drie. (Zie de kaders en Figuur 1 voor de relevante details en verzamelde gegevens van de vier rankings.)

Bekendheid en reputatie

De bekendheid en reputatie van een universiteit hangen natuurlijk af van vele factoren. Twee daarvan zijn relatief

makkelijk te bepalen: omvang en leeftijd. Immers, hoe ouder een instelling, hoe meer lichtingen alumni (staf en student). En hoe groter een instelling, hoe meer alumni per lichting. Naast de netwerken van de zittende studenten en staf, zullen die van de alumni, als potentiële ambassadeurs, een belangrijke bijdrage leveren aan de naamsbekendheid in binnen- en buitenland. Een derde makkelijk te achterhalen factor die vermoedelijk een belangrijke rol speelt, is zichtbaarheid in de bekendste citatiedatabases en/of de relatieve citatie-impact.

Op basis van de methodologische verschillen tussen de vier rankings hebben we de volgende hypothesen geformuleerd:

1. De positie van universiteiten is in de ARWU meer gecorreleerd met grootte dan in de THE en QS.
2. De positie van universiteiten in de THE en QS is als gevolg van de zwaarwegende reputatiescores meer gecorreleerd met de leeftijd van een instelling dan in de ARWU.
3. De op de citatiedatabases gebaseerde onderzoekindicatoren uit de drie buitenlandse rankings zijn gecorreleerd met ten minste één van de geselecteerde Leidse indicatoren.

De hypothesen 1 en 3 blijken we, niet onverwacht, grotendeels te kunnen bevestigen. Maar tot onze eigen verrassing lijken oudere universiteiten ook in de ARWU een duidelijke voorsprong te hebben, ondanks het ontbreken van een reputatiescore. Het is een publiek geheim dat rankings als de THE en QS, die deels afhankelijk zijn van door de instellingen zelf te leveren gegevens, leiden tot strategisch gedrag (*‘gaming’*). Daarnaast zijn er vraagtekens geplaatst bij de representativiteit en zeggenschap van beide reputatiesurveys. Omdat deze twee kritiekpunten niet van toepassing zijn

op de ARWU en Leiden Ranking, achten velen, onder wie ikzelf, die betrouwbaarder en veelzeggender. Maar ons aanvankelijke beeld van 'willekeur en onvergelykbaarheid' van de THE en QS rankings hebben we toch moeten bijstellen. Beider reputatie-indicatoren vertonen meer overeenkomst dan verwacht en hun onderzoeksindicatoren correleren bovendien met die van ARWU en Leiden. De vier rankings blijken deels complementaire informatie te bieden, waarbij THE en QS vooral meerwaarde lijken te bieden voor universiteiten met een focus op *social sciences* en *humanities*.

Alternatieve indicatoren

Door alle vier rankings aan elkaar te koppelen, beschikken we over vier alternatieve indicatoren voor relatieve omvang van de instellingen die in alle vier Top-300-lijsten voorkomen, en wel die op basis van:

1. het aantal Web of Science-publicaties uit de Leiden Ranking (*LEI P* of *LEI Rnk P*);
2. de uit de *Per Capita Performance* (zie Kader 2, Fig. 1b) berekende stafomvang volgens de ARWU (*ARWU Size* of *ARWU Size Rnk*);
3. het aantal studenten (in formatieplaatsen) gepubliceerd door THE (*THE Fte Studs* of *THE Rnk Fte Studs*);
4. de uit studenten en student-stafratio berekende stafomvang uit de THE (*THE FTEStaff* of *THE Rnk FTEStaff*).

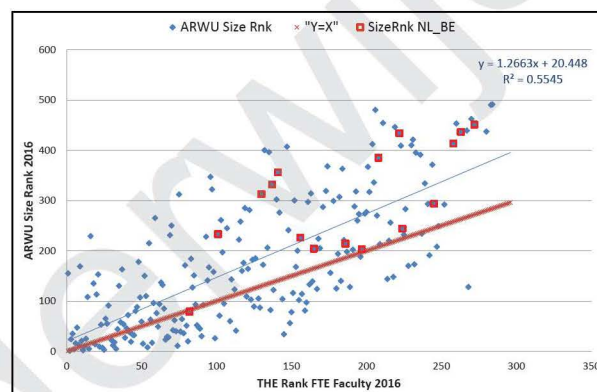
Alvorens de drie geformuleerde hypothesen te toetsen, hebben we eerst bepaald in hoeverre deze alternatieve omvangsindicatoren met elkaar corresponderen. Daartoe zijn de paarsgewijze Spearman-rangcorrelatie coëfficiënten (r_s) bepaald voor de betreffende deelrangnummers²:

Spearman rank coefficients between Size rank indicators Top-300				
	LEI Rnk P	ARWU Size Rnk	THE Rnk FTE Studs	THE Rnk FTE Staff
LEI Rnk P	1			
ARWU Size Rnk	0,66	1		
THE Rnk FTE Studs	0,52	0,51	1	
THE Rnk FTE Staff	0,64	0,74	0,56	1

Figuur 1 Spearman-rangcorrelatie coëfficiënten

Het aantal universiteiten dat in alle vier Top-300-lijsten voorkomt bedraagt 159 (waaronder 11 Nederlandse en 5 Belgische), dus de corresponderende aantallen in de paarsgewijze vergelijkingen zijn groter. Alle omvang indicatoren blijken onderling significant positief gecorreleerd ($p < 0,0002$ bij een $r_s \geq 0,394$ en $n = 60$). De correlatie is het sterkst tussen die volgens ARWU en THE staf, geplot in Figuur 2.

Gegeven de significante correlaties, lijkt het gerechtvaardigd een *size*-indicator uit de ene ranking ook te gebruiken voor een andere. Voor de QS-ranking, die zelf geen informatie publiceert over de omvang van de instellingen (anders dan een indeling in vier grootteklassen) hebben we daarom die uit de ARWU gebruikt als (continue) *size*-indicator.

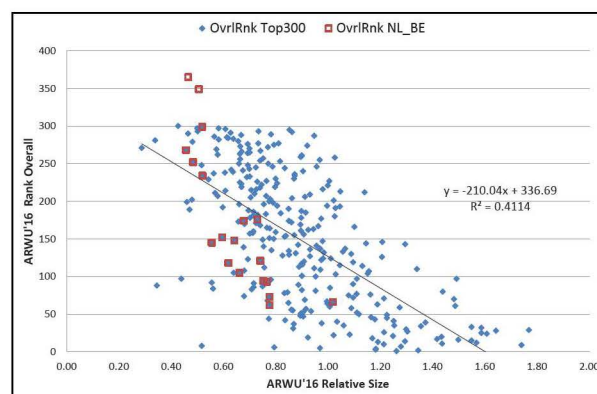


Figuur 2 Rangnummers naar grootte van de Top-300 universiteiten volgens ARWU uitgezet tegen die volgens THE.

Hypothese 1: de positie van universiteiten is in de ARWU meer gecorreleerd met grootte dan in de THE en de QS.

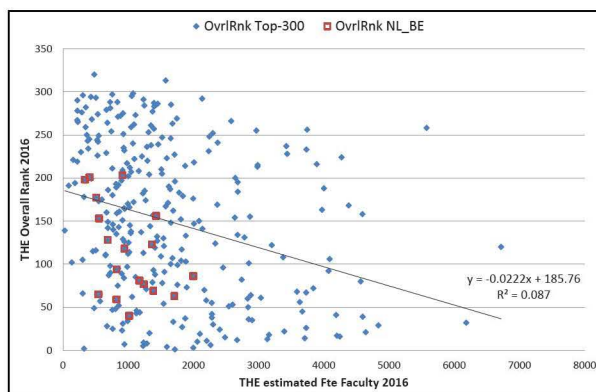
De snelste manier om het effect van grootte op de *overall* (Ovrl) rangnummers te bepalen is door deze rechtstreeks uit te zetten tegen een van de (relatieve) omvangsindicatoren, als in de volgende 3 figuren (Nederlandse en Belgische universiteiten met rode vierkantjes gemarkeerd).

Voor de ARWU levert een plot tegen relatieve omvang een significant negatieve correlatie op (ofwel: 'bigger is better' ~ lager *overall* rangnummer):



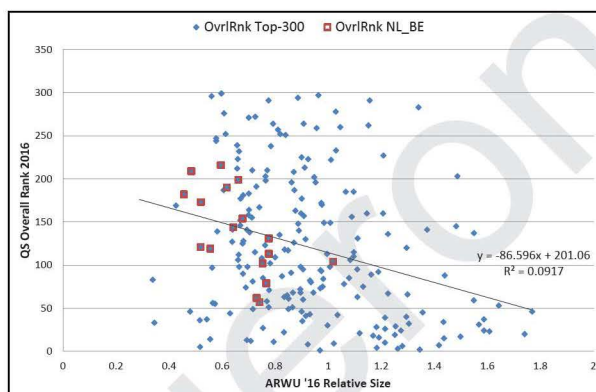
Figuur 3 Overall rangnummers Top-300-universiteiten in de ARWU uitgezet tegen relatieve omvang staf

Voor de **THE** blijkt het verband met stafomvang in formatieplaatsen veel minder sterk:



Figuur 4 Overall rangnummers Top-300 universiteiten in de THE uitgezet tegen omvang staf.

Hetzelfde geldt voor de QS:



Figuur 5 Overall rangnummers Top-300-universiteiten in de QS, uitgezet tegen omvang staf als afgeleid uit de ARWU.

Daarmee lijkt hypothese 1 bevestigd: de Ovrl positie in de ARWU wordt in niet onbelangrijke mate (41 procent verklaarde variantie) bepaald door de omvang van de instelling. Hoe groter, hoe hoger in de Top-300. Voor de THE en QS lijkt dat veel minder het geval.

Hypothese 2: de positie van de universiteiten zijn in THE en QS meer dan in de ARWU of Leiden Ranking gecorreleerd met leeftijd.

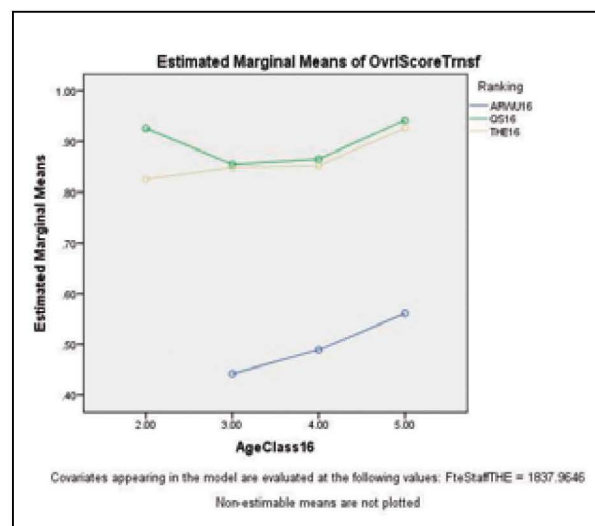
Bij de eenvoudigste statistische toets van enkel het effect van leeftijd, kan het beeld vertroebeld raken door het gelijktijdig effect van grootte en door verschillen tussen de rankings. Daarom hebben we hypothese 2 met een verfijnder statistisch model getoetst, een zogenaamde 2-weg covariantieanalyse⁴. Dat heeft het volgende resultaat opgeleverd:

Tests of Between-Subjects Effects
Dependent Variable: OvrlScoreTrnsf

Source	Type III Sum of Squares	Df	Mean Square	F	Sig.
Corrected Model	19.990 ^a	11	1.817	75.152	.000
Intercept	16.904	1	16.904	699.061	.000
FteStaf-fTHE	.616	1	.616	25.459	.000
Ranking	7.154	2	3.577	147.933	.000
AgeClass	.716	3	.239	9.874	.000
Ranking * AgeClass	.014	5	.003	.118	.988
Error	15.404	637	.024		
Total	454.763	649			
Corrected Total	35.394	648			

^a R Squared = .565 (Adjusted R Squared = .557)

Het effect van *AgeClass* is dus statistisch significant en positief – hoe ouder, hoe hoger in de ranking – maar anders dan verwacht is dit effect vergelijkbaar voor de drie rankings, zoals blijkt uit de niet-significante interactie tussen ranking en *AgeClass* en de volgende figuur:



Hoewel niet significant anders, lijkt het leeftijdseffect in de ARWU tot onze verrassing juist sterker dan in de THE en QS. Opgemerkt zij dat de ogenschijnlijke afwijking van de allerjongste universiteiten in de QS en THE een vertekening is als gevolg van de kleine *sample size*: in beide Top-300-lijsten zit slechts één universiteit in leeftijdsklasse 2 (10-25 jaar).

Overeenkomsten en verschillen in methodologie

De oudste en bekendste mondiale ranking is de *Academic Ranking of World Universities* (ARWU), in 2003 voor het eerst gepubliceerd door de Shanghai Jiao Tong universiteit. In tegenstelling tot de THE en QS, is deze geheel gebaseerd op door de rankers zelf uit openbare bronnen vergaarde kwantitatieve gegevens (Nobelprijzen, Fields Medals, publicaties in de Web of Science en in de toptijdschriften *Nature* of *Science*, *HighlyCited authors*). Er wordt voor deze ranking geen *survey* uitgevoerd en reputatie speelt hierin dus geen (directe) rol. Om die reden wordt de ARWU in het algemeen beschouwd als de meest 'objectieve' en stabiele ranking. Daar staat tegenover dat de ARWU nauwelijks rekening houdt met verschillen in grootte: hoe meer prijzen, (top) publicaties of veelgeciteerden, hoe hoger de totaalscore en -rang⁶.

Geheel anders is dat bij de THE- en QS-rankings die beide bij de deelnemende universiteiten zelf vragen om opgaaf van student- en stafaantallen. Die gegevens gebruiken ze onder meer om andere indicatoren voor omvang te corrigeren – reden waarom beide vermoedelijk niet of minder grootteafhankelijk zijn. Wel leunen ze dus zwaar op reputatiesurveys.

Deze rankings maken alle drie gebruik van gegevens ontleend aan een van de twee belangrijkste citatiedatabases: Web of Science (van de firma Clarivate Analytics, voorheen Thomson Reuters) dan wel *Scopus* (Elsevier). Die gegevens, ingedeeld in de 'dimensie' onderzoek, tellen voor 20 procent (QS), 38,5 procent (THE) of 66 procent (ARWU) mee in de totaalscore. Het restant wordt bepaald door indicatoren in de dimensies onderwijs (alle drie), internationalisering (QS en THE) en

industrie (QS en THE). Zie Figuur 1 voor een overzicht van alle dimensies en indicatoren in deze drie rankings. Gegeven de relatief slechte dekking van de social sciences en de humanities in Web of Science en Scopus, mag je aannemen dat rankings minder zeggingskracht hebben voor instellingen met een focus op deze disciplines naarmate ze zwaarder leunen op indicatoren ontleend aan Web of Science of Scopus.

De Leiden Ranking, als vierde geselecteerd, wijkt sterk af van de andere drie vanwege de bewuste beperking tot uitsluitend aan Web of Science ontleende indicatoren. Eigenlijk is het geen ranking maar een online tool om meer dan 800 universiteiten met voldoende publicaties in de Web of Science op een aantal zelf te kiezen indicatoren te vergelijken. De onderliggende CWTS-methodologie wordt breed erkend als degelijk en betrouwbaar, reden waarom deze ranking is meegenomen als 'benchmark' voor een aantal onderzoekindicatoren van de andere drie. Uit de Leiden Ranking hebben we de volgende drie indicatoren geselecteerd:

1. het aantal publicaties in de Web of Science (P, gesommeerd over een vijfjaarsperiode, zonder 'fractional counting'; als een proxy voor omvang van het onderzoek)
2. de voor discipline, jaar en soort publicatie genormaliseerde citatie-impact score (MNCS)
3. het percentage publicaties dat behoort (na vergelijkbare normalisatie) tot de 10 procent meest geciteerde ter wereld (PP_Top10%).

De laatste twee zijn te beschouwen als schattingen van academische impact van het onderzoek (wel aangeduid als de CWTS 'crown indicators').

Pairwise Spearman rank coefficients between Research indicators Top-300

	LEI P	LEI MNCS	LEI PP-top10%	ARWU Awrd	ARWU HiCi	ARWU N&S	ARWU Pub	THE Res	THE Cit	QS Acad	QS CitFac
LEI P	1										
LEI MNCS	0.42	1									
LEI PPtop10	0.42	0.93	1								
ARWU Awrd	0.30 ^{ns}	0.45	0.43	1							
ARWU HiCi	0.51	0.40	0.43	0.33	1						
ARWU N&S	0.51	0.52 ^{ns}	0.53 ^{ns}	0.47	0.47	1					
ARWU Pub	0.92	0.17	0.18	0.36	0.39	0.34	1				
THE Res	0.55	0.33	0.32	0.56	0.32	0.38	0.50	1			
THE Cit	0.33	0.76	0.76	0.17 ^{ns}	0.48	0.43	-0.03	0.06	1 ^{ns}		
QS Acad	0.63	0.37	0.35	0.53	0.28 ^{ns}	0.40	0.60	0.72	0.15 ^{ns}	1	
QS CitFac	0.17 ^{ns}	0.37	0.41	0.45	0.26	0.35	0.06 ^{ns}	0.39	0.21 ^{ns}	0.26 ^{ns}	1

Daarmee lijkt hypothese 2 gedeeltelijk bevestigd: hoe ouder de universiteit, hoe beter de positie (ofwel: hoe hoger de totaalscore, des te lager het rangnummer) in alle drie rankings. Maar voor de ARWU gaat dit minstens zoveel op als voor de THE en QS.

Hypothese 3: op Web of Science of Scopus gebaseerde onderzoekindicatoren van ARWU, THE en QS zijn gecorreleerd met een van de Leidse Ranking-indicatoren.

Om hypothese 3 te toetsen hebben we de paarsgewijze Spearman-rangcorrelatie coëfficiënten (r_s) bepaald voor de volgende elf deelrangnummers, waarvan acht aan de Web of Science of Scopus ontleend (in groen):

Net als voor de eerder besproken size-indicatoren, zijn alle r_s waarden > 0.39 in elk geval statistisch significant⁵ ($n > 60$); in de tabel hebben we alle met een $r_s < 0.30$ als niet significant (ns) gemarkeerd..

Van de acht op Web of Science of Scopus gebaseerde indicatoren zijn er vier niet genormaliseerd voor omvang (in groen): LEI P (ofwel: het aantal Web of Science-publicaties) en alle drie ARWU-Web of Science-indicatoren (*Highly Cited Authors* HiCi, *Nature & Science publications* N&S, Web of Science-publicaties Pub). Slechts één van deze vier, ARWU Pub, blijkt niet significant gecorreleerd met de twee Leidse *crown indicators* (de *Mean Normalised Citation Score* MNCS en het *Percentage top10% most cited publications*, PPtop10%, die beide niet grootteafhankelijk zijn) maar wel, als te verwachten, heel sterk met LEI P.

De significant positieve correlaties van beide Leidse *crown indicators* met LEI P, ARWU HiCi en ARWU N&S doen vermoeden dat er kennelijk toch voldoende 'kritische massa' nodig is om een hoge (absolute en relatieve) citatie-impact te bereiken.

De vier voor omvang genormaliseerde Web of Science- of Scopus-indicatoren (in vet groen) blijken alle onderling significant positief gecorreleerd op twee citatie-indicatoren na: **QS CitFac en THE Cit**.

Dat laatste valt toe te schrijven aan verschillende onderliggende normalisaties. Waar QS de genormaliseerde citaties (*Scopus citations*, 'field normalised') deelt door de formatieplaatsen van de academische staf, past THE (Scopus citations, 'field normalised + country-adjusted') daarnaast ook nog een (niet-transparante) regionale normalisatie toe.

Dat de andere wel onderling correleren geeft enig vertrouwen in de betrouwbaarheid en vergelijkbaarheid van de onderliggende dataverzameling en bewerkingen.

De wellicht verrassendste significant positieve correlaties in bovenstaande tabel zijn die van de drie niet (uitsluitend) op Web of Science of Scopus gebaseerde indicatoren:

- tussen ARWU Awrd enerzijds en THE Res en QS Acad anderzijds;
- tussen THE Res enerzijds en QS Acad, en in iets mindere mate ook LEI P en ARWU Awrd, anderzijds;
- tussen QS Acad enerzijds en LEI P, en in iets mindere mate ook ARWU Pub en ARWU Awrd, anderzijds.

Een mogelijke verklaring voor a) is dat toekenning van Nobelprijzen en *Fields Medals* de reputatie van een instelling versterkt. Dat komt vervolgens tot uitdrukking in de reputatiedeelcores van zowel THE als QS.

Wat b) betreft: THE Res omvat de hele *research pillar*: 18 procent *reputation*, 6 procent *income* en 6 procent *productivity*; de reputatiedeelcore is hierin dus dominant. QS Acad betreft geheel de reputatie onder *peers*. Hoewel de THE survey aanzienlijk betrouwbaarder lijkt dan die van QS, leveren beide kennelijk toch deels vergelijkbare uitkomsten op en

Verzamelde en afgeleide gegevens

Allereerst hebben we van de betreffende websites de beschikbare gegevens gedownload voor de *Overall* (Ovrl) Top-300 van beste universiteiten in de laatste (2016) edities van alle vier rankings. Naast de Ovrl rangnummers en totaalscores (in ARWU, QS en THE omgezet naar een 0-100 schaal met score 100 voor de nummer 1), zijn dat ook de beschikbare deelscores op de onderliggende indicatoren en bijbehorende deelrangnummers.

Figuur 1 geeft een overzicht van alle deelindicatoren in de drie 'buitenlandse' rankings en hun gewichten. Vanaf rangnummer 101 (ARWU), 201 (THE) of 401 (QS) worden de instellingen gegroepeerd in groepen van 40-50 waarbinnen de onderlinge verschillen statistisch te klein zijn voor toekenning van *separate* ranks en publiceren ARWU en THE ook geen totaalscores meer. Maar die zijn eenvoudig te reconstrueren met behulp van de

wel gepubliceerde deelscores en gewichten. Daarmee konden we alsnog gedetailleerde Ovrl rangnummers berekenen.

De THE publiceert enkel deelscores op de vijf zogenaamde 'pillars' (de vijf dimensies weergegeven in Figuur 1a), waartoe we ons hier dan ook beperkt hebben. Behalve de vijf pillarscores, publiceert de THE op haar website ook de studentaantallen en student-stafratio's van elke universiteit. Daaruit kunnen we ook de stafaantallen eenvoudig afleiden.

Evenzo kunnen we uit de *Per Capita Performance*-deelscore van de ARWU een indicator voor relatieve omvang van de staf berekenen.

QS publiceert op haar website ook de leeftijds- (1: < 10, 2: 10-25, 3: 25-50, 4: 50-100, 5: >100 jaar) en grootteklassen (XL: > 30000, L: 20-30000, M: 5-12000, S: <5000 studenten) waarin een instelling valt.

Figuur 1. Deelindicatoren en gewichten in de drie buitenlandse rankings

Criteria	Indicators	Code	Weight
Quality of Education	Alumni of an institution winning Nobel Prizes and Fields Medals	Alum	10%
Quality of Faculty	Staff of an institution winning Nobel Prizes and Fields Medals	Award	20%
	Highly cited researchers in 21 broad subject categories	HCI	20%
	Papers published in Nature and Science*	N&S	20%
Research Output	Papers indexed in Science Citation Index-expanded and Social Science Citation Index	PUB	20%
Per Capita Performance	Per capita academic performance of an institution	PCP	10%
Total			100%

* For institutions specialized in humanities and social sciences such as London School of Economics, N&S is not considered, and the weight of HCI is increased to other indicators.

b) ARWU, bron: www.shanghairanking.com/ARWU-Methodology-2016.html

INDICATORS - WEIGHTINGS & DEFINITIONS					
Academic Reputation	Academic Reputation	Academic Reputation	Academic Reputation	Academic Reputation	Academic Reputation
Employment Reputation	Employment Reputation	Employment Reputation	Employment Reputation	Employment Reputation	Employment Reputation
International Faculty	International Faculty	International Faculty	International Faculty	International Faculty	International Faculty
International Students	International Students	International Students	International Students	International Students	International Students

a) THE, bron: www.timeshighereducation.com/world-university-rankings/methodology-world-university-rankings-2016-2017

c) QS, bron: <https://www.topuniversities.com/qs-world-university-rankings/methodology>

lijken beide positief beïnvloed te worden door zichtbaarheid in de Web of Science en/of door 'kritische massa' en toekenning van Nobelprijzen en Fields Medals.

Correlatie c) ten slotte bevestigt genoemd vermoeden dat een grotere zichtbaarheid, zowel in Scopus als Web of Science, en/of 'kritische massa' ook bevorderlijk is voor de reputatie onder QS-peers, evenals toekenning van Nobelprijzen en Fields Medals.

Grotendeels bevestigd

Onze hypothesen zijn, kortom, grotendeels bevestigd. Samengevat: omvang van de instelling speelt duidelijk een grotere rol in de ARWU dan in THE en QS. Maar kritische massa en zichtbaarheid in Web of Science en Scopus of via Nobelprijzen en Fields Medals lijken indirect ook in de QS en THE door te werken via de reputatiescores. En ondanks het ontbreken van een reputatiescore in de ARWU, blijken ook daarin oudere universiteiten duidelijk in het voordeel.

Al met al hebben we ons aanvankelijke beeld van 'willekeur en onvergelykbaarheid' van de drie buitenlandse rankings toch enigszins bij moeten stellen. De onderzoeks- en reputatieindicatoren vertonen meer overeenkomst dan we verwachtten. Daarnaast lijkt de conclusie gerechtvaardigd dat de vier rankings deels complementaire informatie bieden. Waar de relevantie van de Leiden Ranking en ARWU zich vooral beperkt tot *research universities* met een focus op bèta- en medische wetenschappen, hebben de THE en QS

potentieel meer zeggingskracht voor universiteiten met een focus op *social sciences* en *humanities*.

Deze aflevering van de datarubriek werd samengesteld door Jules van Rooij, adviseur onderzoeksbeleid en institutional research aan de Rijksuniversiteit Groningen

Noten

- 1 Ontwikkeld door het Leidse Centrum voor Wetenschap & Technologie Studies (CWTS), zie Waltman et al., 2012, *The Leiden Ranking 2011/2012: Data Collection, Indicators, and Interpretation*, J Am Soc Inf Sc & Tech, 63(12): 2419–243
- 2 Berekening van de paarsgewijze correlatiecoëfficiënten tussen de deelscores levert een vrijwel identieke uitkomst op.
- 3 Het ontbreken van een grootte-effect op de totaalscore in de QS-ranking (genormaliseerd door middel van een arcsinustransformatie) wordt ook bevestigd door een ANOVA met QS sizeclass als onafhankelijke variabele: $F_{3,296} = 0.337$, $p = 0.80$.
- 4 Daartoe is gebruikgemaakt van een tweewegcovariantieanalyse ter gelijktijdige toetsing van de effecten van ranking en de aan QS ontleende discrete leeftijdsindicator, Age-Class, als onafhankelijke variabelen en de aan THE ontleende FTEStaff als covariabele, op de (met behulp van een arcsinustransformatie genormaliseerde) totaalscores in de drie rankings. Voorwaarde voor zo'n parametrische toets is dat de data normaal verdeeld zijn. Dat gaat niet op voor rangnummers, noch voor scores op een 0-100 schaal. Na een arcsinustransformatie ($\sin^{-1} \sqrt{\text{score}/100}$) bleken de totaalscores van de drie buitenlandse rankings wel normaal verdeeld.
- 5 Feitelijk dient nog gecorrigeerd te worden voor het grote aantal paarsgewijze vergelijkingen; het gaat ons hier evenwel niet om de werkelijke statistische waarschijnlijkheid, maar om een ruwe indicatie.
- 6 De enige van de zes ARWU-indicatoren die wel gecorrigeerd wordt voor omvang, is de Per Capita Performance, maar die telt slechts voor 10 procent mee in de totaalscore.